

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
2 June 2005 (02.06.2005)

PCT

(10) International Publication Number
WO 2005/050615 A1

(51) International Patent Classification⁷: **G10G 3/04**,
G10L 11/04, G11B 31/00, G06F 17/00, G10H 7/00

(74) Agent: **ALBAN TAY MAHTANI & DE SILVA**; 39
Robinson Road, #07-01, Robinson Point, Singapore
068911 (SG).

(21) International Application Number:
PCT/SG2003/000276

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE,
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT,
RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date:
21 November 2003 (21.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (*for all designated States except US*): **AGENCY
FOR SCIENCE, TECHNOLOGY AND RESEARCH**
[SG/SG]; 20 Biopolis Way, #07-01, Centros, Singapore
138668 (SG).

(84) Designated States (*regional*): ARIPO patent (BW, GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,
SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventor; and

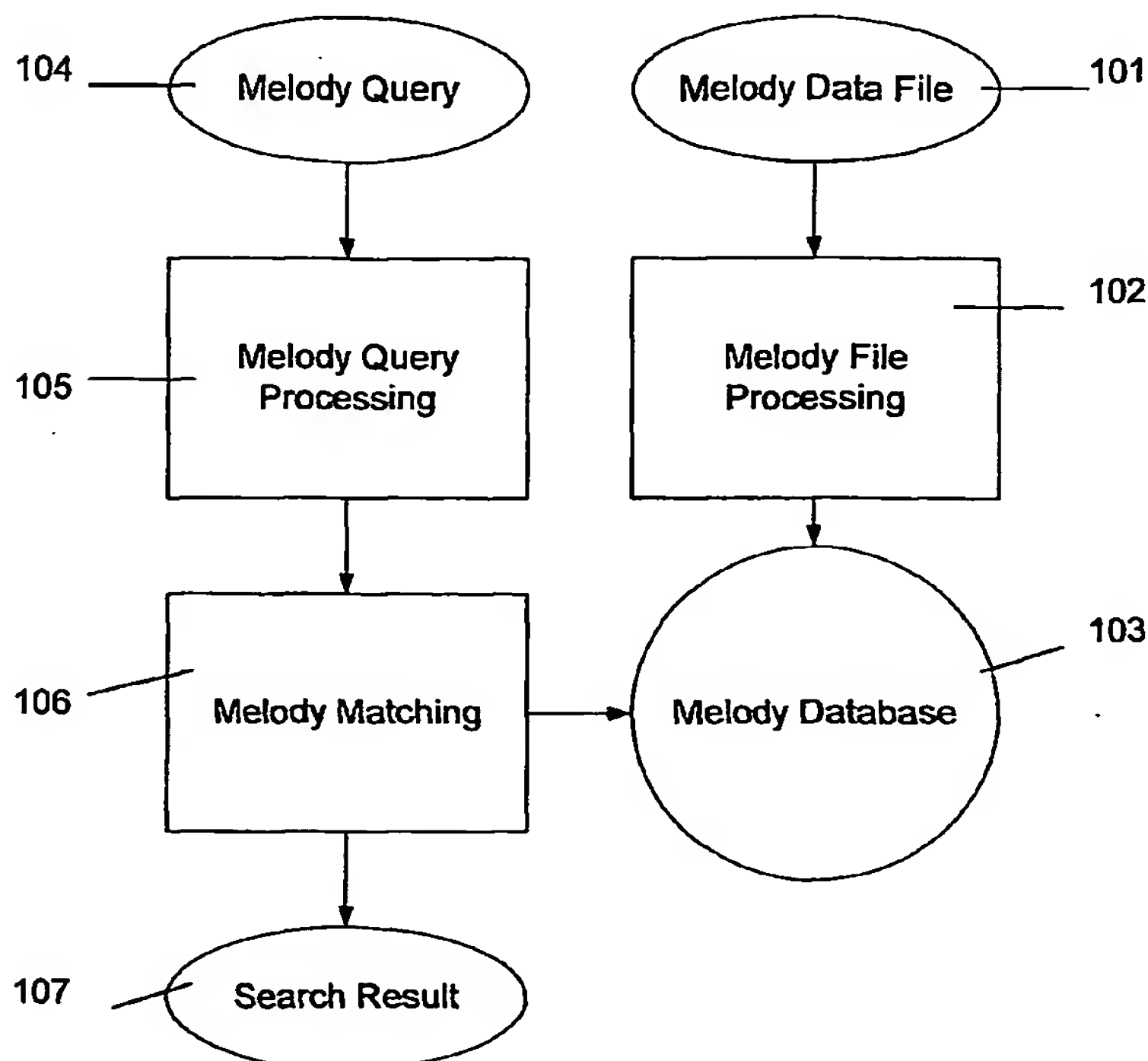
(75) Inventor/Applicant (*for US only*): **ZHU, Yongwei**
[SG/SG]; BLK 320C Anchorvale Drive, #13-124, Singa-
pore 543320 (SG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv)) for US only

[Continued on next page]

(54) Title: **METHOD AND APPARATUS FOR MELODY REPRESENTATION AND MATCHING FOR MUSIC RETRIEVAL**



(57) Abstract: This invention discloses a method for melody representation and matching able to accommodate pitch and speed variations in the query input. The melody is represented by a sequence of data points, which is invariant to the speed or tempo of the melody. For the melody representation, the hummed query is converted to a pitch time series. The pitch time series is then approximated by a sequence of line segments. The line segment sequence in time domain is then mapped into a sequence of points in a value-run domain. The sequence of points is invariant to the time or speed in the original time series. In a data point sequence matching technique, the query data sequence is aligned with the target data sequence in a database. This alignment is done based on important anchor points in the data sequences that can tolerate value variation (pitch and key inaccuracy in the hummed query) and it also helps determine the probable matching candidates from all the subsequences of the target data sequences. The similarity between the query data sequence with the aligned candidate data subsequence is computed

using a melodic similarity metric, which is based on melody aligning.



Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.